

## برآورد مجموع در آمارگیری‌های سه‌چارچوبی

طیبه چگینی\* و حمیدرضا نواب‌پور

دانشگاه علامه طباطبایی

چکیده: در آمارگیری از جامعه‌های متناهی معمولاً یک نمونه‌ی احتمالاتی از یک چارچوب نمونه‌گیری که واحدهای جامعه‌ی هدف را پوشش می‌دهد انتخاب می‌شود. در برخی موارد، چارچوبی که در دسترس است به‌صورت کامل جامعه‌ی هدف را پوشش نمی‌دهد. لذا آماره‌های آمارگیری با استفاده از این چارچوب‌ها اغلب اریب هستند. از طرفی روزآمدسازی چارچوبی که در دسترس است نیاز به صرف هزینه‌های زیاد دارد. از آمارگیری‌های چندچارچوبی با هدف افزایش پوشش جامعه و بهبود میزان اریبی و کاهش هزینه‌ها استفاده می‌شود. در آمارگیری‌های چندچارچوبی، نمونه‌ها به‌صورت مستقل از چندچارچوب که پوشش مناسبی از جامعه‌ی هدف ارائه می‌کنند و هم‌پوشانی دارند انتخاب می‌شوند. چندین برآوردگر برای مجموع جامعه‌ای در آمارگیری‌های چندچارچوبی مطرح شده است. ما در این مقاله ضمن معرفی این برآوردگرها، در یک مطالعه‌ی شبیه‌سازی کارایی برآوردگرهای مختلف را با هم مقایسه می‌کنیم.

واژگان کلیدی: چارچوب نمونه‌گیری؛ آمارگیری‌های چندچارچوبی؛ برآوردگر هارتلی؛ برآوردگر فولر - بورمیستر؛ برآوردگر ماکسیمم درست‌نمایی‌نما.

### ۱- مقدمه

در نمونه‌گیری به روش کلاسیک از جامعه‌ی متناهی، یک نمونه‌ی احتمالاتی از یک چارچوب نمونه‌گیری که شامل واحدهای جامعه‌ی هدف است، انتخاب می‌شود. ولی از آنجایی که معمولاً در عمل پوشش هیچ چارچوبی کامل نیست، راه مناسب‌تر آن است که نمونه‌های مستقلی از  $Q$  چارچوب نمونه‌گیری متفاوت، که اجتماع آن‌ها جامعه‌ی مورد نظر

---

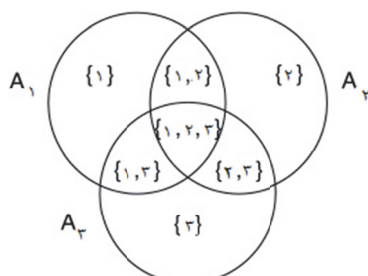
\* نویسنده‌ی عهده‌دار مکاتبات  
دریافت: ۱۳۹۲/۱۰/۸، پذیرش: ۱۳۹۳/۹/۳.

را پوشش می‌دهد، انتخاب شود و سپس اطلاعات به‌دست آمده از نمونه‌ها برای برآورد کمیت‌های جامعه ترکیب شوند.

برای مثال یک آمارگیری در جامعه‌ی آمارشناسان ایالت‌های متحده را در نظر بگیرید. در این آمارگیری یکی از چارچوب‌های نمونه‌گیری می‌تواند فهرست عضوهای مستقیم انجمن امریکایی آمار (ASA) باشد که یک نمونه‌ی احتمالاتی از افرادی که عضو فهرست اعضای این انجمن هستند انتخاب می‌شود. بیش‌ترین افراد در این نمونه احتمالاً آمارشناس بوده و بنا بر این چارچوب مورد نظر مقرون به صرفه خواهد بود. با این حال بسیاری از آمارشناسان در ایالت‌های متحده عضو ASA نیستند، در نتیجه فهرست اعضای این انجمن، جامعه‌ی آمارشناسان ایالت‌های متحده را به‌طور کامل پوشش نمی‌دهد. پوشش این چارچوب را می‌توان با یک چارچوب دوم مانند فهرست عضوهای مستقیم مؤسسه‌ی آمار ریاضی (IMS) بهبود بخشید. به این ترتیب یک نمونه‌ی احتمالاتی از چارچوب  $A_1$  (عضوهای مستقیم ASA) و مستقل از یک نمونه‌ی احتمالاتی از چارچوب  $A_2$  (عضوهای مستقیم IMS) در نظر گرفت. بسیاری از آمارشناسان به هر دو سازمان تعلق دارند، بنا بر این دو چارچوب هم‌پوشانی دارند. چارچوب  $A_1$  و  $A_2$  با هم، پوشش بهتری برای جامعه فراهم می‌کنند اما هنوز تمام آمارشناسان را شامل نمی‌شوند. در این حالت، یک آمارگیری سه‌چارچوبی می‌تواند مورد استفاده قرار گیرد. برای همین منظور چارچوب  $A_3$  ممکن است یک چارچوب از تمام جامعه‌ی بزرگسالان باشد.

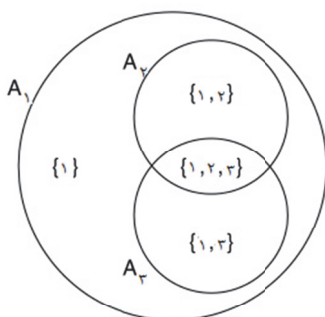
ممکن است این پرسش مطرح شود که چرا از همان ابتدا از چارچوب  $A_3$  که شامل تمام جامعه است یک نمونه گرفته نمی‌شود؟ در واقع به دلیل هزینه‌ی زیاد آمارگیری از این چارچوب، از دو چارچوب ناقص دیگر استفاده شده است.

هاینس و پولاک استفاده از آمارگیری‌های چندچارچوبی را برای برآورد اندازه‌ی جامعه مطرح کردند و به بررسی تناظر بین آمارگیری‌های چندچارچوبی و روش‌های گیر و بازگیر پرداختند [۳]. آن‌ها اطلاع‌های موجود از چارچوب‌های فهرستی ناقص و یک چارچوب ناحیه‌ای را برای برآورد تعداد عقاب‌ها در یک ناحیه ترکیب کردند. ایچان و دنیس طرح چندچارچوبی را برای نمونه‌گیری از جامعه‌ی بی‌خانمان‌ها استفاده کردند که چارچوب‌های مورد استفاده در این طرح  $A_1$  پناهگاه‌های بی‌خانمان‌ها،  $A_2$  نوانخانه‌ها و  $A_3$  مکان‌های خیابانی بودند [۷]. ساختار این طرح در شکل ۱ به تصویر کشیده شده است [۱۱].



شکل ۱- مثالی از یک آمارگیری سه‌چارچوبی

در جامعه‌های کمیاب نیز اغلب، استفاده از آمارگیری‌های چندچارچوبی می‌تواند نمونه‌های کاراتری به‌دست دهد. برای مثال، در مطالعه‌ی مشخصه‌ای از افراد مبتلا به آسم یکی از چارچوب‌ها می‌تواند چارچوبی باشد که در سرشماری سلامت از کل جامعه به‌دست آمده در حالی که می‌توان از دیگر چارچوب‌ها نظیر فهرست افراد در کلینیک‌های مربوط به یک بیماری خاص، مراکز درمانی و بیمارستان‌ها استفاده کرد. با آمارگیری‌های چندچارچوبی انتظار می‌رود که برآوردهای دقیق‌تری از افراد مبتلا به بیماری آسم به‌دست آید، چرا که با سه چارچوب یا بیش‌تر، پوشش کامل‌تری از افراد مبتلا به بیماری آسم به‌دست خواهد آمد. ساختار این طرح در شکل ۲ نشان داده شده است. چارچوب  $A_1$  گرچه چارچوبی از کل واحدهای جامعه است اما ممکن است پوشش  $A_2$  و  $A_3$  در برخی بخش‌های جامعه کامل‌تر باشد.



شکل ۲- مثال دیگری از یک آمارگیری سه‌چارچوبی

اغلب یکی از هدف‌ها در تحلیل داده‌های آمارگیری‌های سه‌چارچوبی برآورد مجموع جامعه‌ای  $Y$  است. در آمارگیری‌های چندچارچوبی، برآوردهای مختلفی برای برآورد

مجموع جامعه‌ای پیش‌نهاد شده است. هارتلی [۵، ۶] و فولر و بورمیستر [۲] به معرفی برآوردگرهایی پرداختند که واریانس را در میان رده‌ای از برآوردگرهای ناریب خطی  $Y$  مینیمم می‌کند. بانکی‌یر [۱] و کالتون و اندرسون [۸] برآوردگرهای تک‌چارچوبی را در مشاهده‌هایی که بر اساس احتمال شمول‌شان در چارچوب‌های مختلف وزن‌دهی می‌شوند، گسترش دادند. اسکینر [۱۶] برآوردگرهای نسبتی چنگک‌زنی را برای زمانی که از هر چارچوب نمونه‌های تصادفی انتخاب شوند، پیش‌نهاد کرد. اسکینر و رائو [۱۷] برآوردگر ماکسیمم درست‌نمایی‌ها را برای  $Y$  در آمارگیری‌های دوچارچوبی با استفاده از طرح‌های پیچیده به دست آوردند. هر چند این برآوردگر برای آمارگیری‌هایی با بیش از دو چارچوب نیز محاسبه می‌شود. لوه‌ر و رائو [۹] کارایی برآوردگرهای مجموع جامعه‌ای را در آمارگیری‌های دوچارچوبی مقایسه کردند. همچنین آن‌ها برآوردگرهای خطی بهین و درست‌نمایی‌ها را برای چارچوب‌های چندگانه (با بیش از دو چارچوب) گسترش دادند. در سال‌های اخیر برآوردگرهای جدیدی نیز برای این آمارگیری‌ها پیش‌نهاد شده‌اند. از جمله مه‌کاتی [۱۴] برآوردگر چندبارگی تک‌چارچوبی را برای این آمارگیری‌ها ارائه داد. لوه‌ر و لو [۱۳] برآوردگرهای جریان ناخالص در آمارگیری دوچارچوبی را مطرح ساختند. رائو و وو [۱۵] نیز روش تجربی درست‌نمایی‌ها را برای استنباط از آمارگیری‌های چندچارچوبی به کار بردند.

در بخش ۲ نمادها برای برآورد مجموع جامعه‌ای در آمارگیری‌های چندچارچوبی معرفی می‌شوند. در بخش ۳ برآوردگرها در آمارگیری‌های چندچارچوبی ارائه می‌شوند. در پایان با انجام یک مطالعه‌ی شبیه‌سازی، به مقایسه‌ی این برآوردگرها در یک آمارگیری سه‌چارچوبی پرداخته می‌شود.

## ۲- نمادها

در آمارگیری‌های چندچارچوبی، نمونه‌های احتمالاتی به صورت مستقل از چارچوب‌های  $A_1, \dots, A_Q$ ،  $Q \geq 2$  انتخاب می‌شوند. فرض کنید  $F = \{1, \dots, Q\}$  مجموعه‌ی اندیس چارچوب‌ها باشد، به طوری که اجتماع این چارچوب‌ها جامعه‌ی مورد نظر را بپوشاند. با

$Q$  چارچوب  $1 - 2^Q$  حوزه‌ی مجزای ممکن وجود دارند که توسط زیرمجموعه‌های  $F$  تعریف می‌شوند. برای  $K \subseteq F$  حوزه‌ی تعریف‌شده توسط  $K$  چنین است:

$$D_K = \left( \bigcap_{i \in K} A_i \right) \cap \left( \bigcap_{j \notin K} A_j^c \right)$$

که  $A_j^c$  مکمل مجموعه‌ی  $A_j$  است. تعداد چارچوب‌های سهم در ساختن حوزه‌ی  $K$  با  $|K|$  نشان داده می‌شود. فرض کنید  $N^{(q)}$  تعداد واحدهای جامعه در چارچوب  $A_q$  و  $N_K$  تعداد واحدهای جامعه در حوزه‌ی تعریف شده توسط  $K$  باشد. از آنجایی که حوزه‌ها متداخل نیستند اندازه‌ی جامعه از رابطه‌ی  $\sum_{K \subseteq F} N_K = N$  محاسبه می‌شود.

دو مثال از آمارگیری سه‌چارچوبی در شکل‌های ۱ و ۲ نشان داده شده‌اند. شکل ۱، یک طرح سه‌چارچوبی را نشان می‌دهد که در آن هر یک از چارچوب‌ها ناقص هستند. حوزه‌های  $\{1\}$ ،  $\{2\}$  و  $\{3\}$  به ترتیب تنها در چارچوب‌های  $A_1$ ،  $A_2$  و  $A_3$ ، حوزه‌های  $\{2,1\}$ ،  $\{3,1\}$  و  $\{3,2\}$  دقیقاً در دو چارچوب ظاهر شده‌اند و حوزه‌ی  $\{3,2,1\}$  اشتراک سه چارچوب است. شکل ۲ نیز یک طرح سه‌چارچوبی است با این تفاوت که چهار حوزه دارد. چارچوب  $A_1$  کامل است اما چارچوب‌های  $A_2$  و  $A_3$  ناقص هستند و با هم تداخل دارند.

مجموع جامعه‌ای  $Y$  به صورت جمع مجموع جامعه‌ای حوزه‌های مجزا نوشته می‌شود:

$$Y = \sum_{K \subseteq F} Y_K$$

که در آن  $Y_K = \sum_{i=1}^N \delta_i(K) y_i$ . اگر واحد مشاهده‌ای  $i$  عضو حوزه‌ی تعریف‌شده توسط  $K$  باشد  $\delta_i(K)$  برابر با یک خواهد بود و در غیر این صورت مقدار صفر می‌گیرد. در حالت خاص، که تعداد واحدهای جامعه در حوزه‌ی  $D_K$  به‌ازای همه‌ی  $i$ ‌ها برابر یک باشد، با در نظر گرفتن  $y_i = 1$  اندازه‌ی حوزه از رابطه‌ی  $N_K = \sum_{i=1}^N \delta_i(K)$  به‌دست می‌آید. مشکلی که در برآورد مجموع جامعه‌ای  $Y$  وجود دارد این است که در هر حوزه مؤلفه‌های  $Y_K$  کم برآورد می‌شوند، به‌همین دلیل از برآورد  $|K|$  در چارچوب‌های نمونه‌گیری مختلف استفاده می‌شود.

فرض کنید  $S_q$  نشان دهنده‌ی نمونه‌ای احتمالاتی از چارچوب  $A_q$  برای  $q = 1, \dots, Q$  باشد. در آمارگیری‌های چندچارچوبی چون نمونه‌ها از  $Q$  چارچوب به‌طور مستقل انتخاب می‌شوند، به تعداد  $|K|$  برآوردگر مستقل از  $Y_K$ ‌ها وجود دارد. برای  $q \in K$ ، برآوردگر مجموع با استفاده از نمونه‌های چارچوب  $A_q$  به‌صورت زیر برآورد می‌شود:

$$(۱) \quad \hat{Y}_K^{(q)} = \sum_{i \in S_q} w_i^{(q)} \delta_i(K) y_i$$

وزن‌های  $w_i^{(q)}$  می‌توانند وارون احتمال شمول یا وزن هایک [۴] باشد. به‌طور مشابه برآوردگر اندازه‌ی حوزه‌ای این چنین تعریف می‌شود:

$$(۲) \quad \hat{N}_K^{(q)} = \sum_{i \in S_q} w_i^{(q)} \delta_i(K).$$

فرض کنید در تمام حالت‌ها واحدها را در هر یک از آمارگیری‌ها می‌توان تشخیص داد. البته این یک فرض قوی است که ممکن است همیشه درست نباشد [۱۰].

### ۳- برآوردگرهای مجموع جامعه‌ای در آمارگیری‌های چندچارچوبی

در این بخش نحوه‌ی محاسبه‌ی برآوردگرهای مجموع جامعه‌ای در آمارگیری‌های چندچارچوبی ارائه شده است. این برآوردگرها عبارت‌اند از: برآوردگرهای خطی بهین، برآوردگر ماکسیمم درست‌نمایی‌نما، برآوردگر تک‌چارچوبی و برآوردگر تعدیل وزنی.

#### ۳-۱- برآوردگرهای خطی بهین

در این بخش ابتدا برآوردگرهای بهین (هارتلی و فولر-بورمیستر) برای طرح‌های آمارگیری عمومی معرفی می‌شوند. حوزه‌های ناخالی با  $K_1, \dots, K_d$  برجسب‌گذاری می‌شوند که البته ممکن است این حوزه‌های ناخالی از ساختار چارچوب‌ها مانند شکل ۱ شناسایی شوند یا زمانی که اندازه‌ی نمونه به حد کافی بزرگ باشد احتمال گزینش یک اندازه‌ی نمونه‌ی صفر

از حوزه‌های ناخالی ناچیز باشد. فرض کنید  $\theta = (\theta_1^T, \dots, \theta_Q^T)^T$  و  $\theta_q = (\theta_{K_1}^{(q)}, \dots, \theta_{K_d}^{(q)})^T$  بردارهایی از ثابت‌ها و مشخص باشند. همچنین برآورد مجموع جامعه‌ای و اندازه‌ی جامعه‌ای در  $q$  آمین چارچوب به‌صورت  $\hat{Y}_q = (\hat{Y}_{K_1}^{(q)}, \dots, \hat{Y}_{K_d}^{(q)})^T$

$\hat{N}_q = (\hat{N}_{K_1}^{(q)}, \dots, \hat{N}_{K_d}^{(q)})^T$  تعریف می‌شوند که اگر حوزه‌ی  $K_i$  بخشی از چارچوب  $A_q$  نباشد یعنی  $q \notin K_i$  آن‌گاه  $\hat{Y}_{K_i}^{(q)} = \hat{N}_{K_i}^{(q)} = 0$  و  $\theta_{K_i}^{(q)} = 0$ . ماتریس قطری  $d \times d$ ،  $\Gamma_q$  طوری تعریف می‌شود که  $i$ امین عنصر قطری‌اش یک است اگر حوزه‌ی  $K_i$  بخشی از چارچوب  $A_q$  باشد ( $q \in K_i$ ) و در غیر این صورت صفر است. نماد  $\mathbf{1}_d$  بردار  $d$  بُعدی از یک‌ها در نظر گرفته می‌شود. بنا بر این برآورد هارتلی  $Y$  برای مقدارهایی از  $\theta$  برابر است با:

$$(3) \quad \hat{Y}_\theta = \sum_{q=1}^Q \theta_q^T \Gamma_q \hat{Y}_q$$

اگر بردار ثابت  $\theta$  در قید زیر صدق کند

$$(4) \quad \sum_{q=1}^Q \Gamma_q \theta_q = \mathbf{1}_d$$

و  $E[\hat{Y}_K^{(q)}] = Y_K$  برای هر حوزه‌ی  $K$  و چارچوب  $A_q$  برقرار باشد، آن‌گاه  $\hat{Y}_\theta$  یک برآوردگر ناریب برای مجموع جامعه‌ای  $Y$  خواهد بود.

برآوردگر هارتلی تعمیم‌یافته از جایگزینی  $\theta_H$  به جای  $\theta$  حاصل می‌شود به طوری که  $\theta_H$  واریانس  $\hat{Y}_\theta$  را تحت قید (۴) مینیمم می‌کند. با تعریف  $C_q$  به عنوان ماتریس کوواریانس  $\hat{Y}_q$  و مینیمم کردن واریانس  $\hat{Y}_\theta$  از رابطه‌ی:

$$V(\hat{Y}_\theta) = \sum_{q=1}^Q \theta_q^T \Gamma_q C_q \Gamma_q \theta_q$$

می‌توان  $\theta_H$  را از حل معادله‌ی زیر به دست آورد:

$$(5) \quad \mathbf{A}\theta = \begin{bmatrix} \mathbf{1}_d^T & \mathbf{1}_d^T \\ 0 & d \end{bmatrix}^T$$

که در آن  $\mathbf{1}_d$  یک بردار  $dQ$  بُعدی ستونی از صفرهاست،  $\mathbf{A}$  یک ماتریس بلوکی  $dQ \times (Q+1)dQ$  بُعدی است به طوری که  $(i, j)$ امین بلوک آن برابر است با:

$$A_{i,j} = \begin{cases} \Gamma_i \left( \sum_{q=1}^Q \Gamma_q \right)^{-1} \left( I - \sum_{q=1}^Q \Gamma_q \right) E_i & i = j \\ \Gamma_i \left( \sum_{q=1}^Q \Gamma_q \right)^{-1} E_j & i \neq j, i \leq Q \\ \Gamma_j & i = Q + 1 \end{cases}$$

همچنین  $I$  ماتریس همانی  $d \times d$  و  $E_i = \Gamma_i C_i \Gamma_i$  است. حال به سراغ براوردگر بهین دیگر یعنی براوردگر فولر-بورمیستر می‌رویم. برای تعمیم براوردگر فولر-بورمیستر، از بردارهای اندازه‌ی حوزه‌های براوردشده‌ی  $\widehat{N}^{(q)}$  استفاده می‌شود. فرض کنید  $\beta = (\beta_1^T, \dots, \beta_Q^T)^T$  یک بردار  $1 \times \sum_{q=1}^Q 2dQ$  از پارامترها باشد و ماتریس  $G_q$  به صورت  $G_q = \text{diag}(\Gamma_q, \Gamma_q)$  تعریف شود. آنگاه براوردگر فولر-بورمیستر برابر خواهد بود با:

$$(۶) \quad \widehat{Y}_{FB}(\beta) = \sum_{q=1}^Q \beta_q^T G_q [\widehat{Y}_q^T \quad \widehat{N}_q^T]^T$$

که  $\beta$  در قید زیر صدق می‌کند:

$$(۷) \quad \sum_{q=1}^Q G_q \beta_q = \begin{bmatrix} 1^T & 0^T \\ 0^T & 1^T \end{bmatrix}$$

با تعریف ماتریس کوواریانس

$$D_q = \text{COV} \begin{bmatrix} \widehat{Y}_q \\ \widehat{N}_q \end{bmatrix} = \begin{bmatrix} C_q & D_{q12} \\ D_{q12}^T & D_{q22} \end{bmatrix}$$

واریانس براوردگر فولر-بورمیستر به صورت زیر در می‌آید:

$$V(\widehat{Y}_{FB}) = \sum_{q=1}^Q \beta_q^T G_q D_q G_q \beta_q$$

مشابه (۵) واریانس براوردگر فولر-بورمیستر نسبت به قید (۷) مینیمم می‌شود زمانی که  $\beta$  در معادله‌ی زیر صدق کند:

$$B\beta = \begin{bmatrix} 1^T & 0^T & 0^T \\ 0^T & 1^T & 0^T \\ 0^T & 0^T & 1^T \end{bmatrix}$$



که در آن بلوک‌های  $2d(Q+1) \times 2dQ$  بُعدی ماتریس  $B$  مشابه ماتریس  $A$  در برآوردگر هارتلی تعریف می‌شوند، با این تفاوت که  $G_i$  جایگزین  $F_i$  خواهد شد [۱۰].

### ۲-۳- برآورد ماکسیمم درست‌نمایی‌ما در نمونه‌گیری‌های تصادفی ساده

یک نمونه‌ی تصادفی ساده به اندازه‌ی  $n^{(q)}$  از چارچوب  $A_q$ ، برای  $q = 1, \dots, Q$  را در نظر بگیرید. فرض کنید  $f = (f^{(1)}, \dots, f^{(Q)})$  بردار  $Q$  بُعدی از کسرهای نمونه‌گیری با  $f^{(q)} = \frac{n^{(q)}}{N^{(q)}}$  باشد. لازم به ذکر است که برای نمونه‌گیری تصادفی ساده وزن نمونه‌گیری واحد  $i$ ام جامعه  $(i = 1, \dots, N^{(q)})$  عکس کسر نمونه‌گیری است. به عبارت دیگر  $w_i^{(q)} = \frac{1}{f^{(q)}}$  بنا بر این  $n_K^{(q)} = \hat{N}_K^{(q)} f^{(q)}$  اندازه‌ی نمونه‌ای از چارچوب  $A_q$  در حوزه‌ی  $K$  است که  $\hat{N}_K^{(q)}$  در (۲) تعریف شده است. در آمارگیری چندچارچوبی می‌توان مجموع جامعه‌ای را به صورت زیر نوشت:

$$Y = \sum_{K \subseteq F} Y_K = N^T \bar{Y}$$

که در آن  $N = (N_{K_1}, \dots, N_{K_d})^T$ ،  $Y = (Y_{K_1}, \dots, Y_{K_d})^T$  و  $\bar{Y}$  بردار متناظر با میانگین حوزه‌ها هستند. لوهر و رائو [۱۰] برآوردهای ماکسیمم درست‌نمایی را برای  $\bar{Y}$  و  $N$  محاسبه کردند به طوری که برآورد ماکسیمم درست‌نمایی‌ما برای  $Y$  برابر است با:

$$(۸) \quad \hat{Y} = \hat{N}^T \hat{\bar{Y}} = \sum_{K \subseteq F} \hat{N}_K \hat{Y}_K$$

برآوردگر  $\bar{y}_K^{(q)} = \sum_{i \in S_q} \delta_i(K) y_i / n_K^{(q)}$  تحت شرایط نظم و با استفاده از قضیه‌ی حد مرکزی، برای نمونه‌های به اندازه‌ی کافی بزرگ، دارای توزیع تقریبی نرمال با میانگین  $\bar{Y}_K$  و واریانس  $\sigma_K^2 / n_K^{(q)}$  است. همچنین برای هر  $q$ ، تحت نمونه‌گیری تصادفی ساده با کسرهای نمونه‌گیری کوچک،  $\{n_K^{(q)}\}$  به طور تقریبی دارای توزیع چندجمله‌ای با اندازه‌ی  $n^{(q)}$  و احتمال موفقیت  $N_K / N^{(q)}$  برای حوزه‌ی  $K$  در چارچوب  $A_q$  است. بنا بر این تابع درست‌نمایی  $(\bar{Y}, N)$  با  $L(\bar{Y}, N) = L_1(\bar{Y}) L_2(N)$  تقریب زده می‌شود که در آن  $C$  مقدار ثابت است.

$$(۹) \quad \log L_{\gamma}(\bar{Y}) = c - \frac{1}{\gamma} \log \sigma_K^2 - \sum_{q=1}^Q \sum_{K:q \in K} \frac{n_K^{(q)} (\bar{y}_K^{(q)} - \bar{y}_K)^2}{\gamma \sigma_K^2}$$

و

$$(۱۰) \quad \begin{aligned} \log L_{\gamma}(N) &= c + \sum_{q=1}^Q \sum_{K:q \in K} n_K^{(q)} \log N_K \\ &= c + \sum_{q=1}^Q \sum_{K:q \in K} \hat{N}_K^{(q)} f^{(q)} \log N_K \end{aligned}$$

با در نظر گرفتن قید

$$(۱۱) \quad N^{(q)} = \sum_{K:q \in K} N_K, \quad q = 1, \dots, Q$$

و با ماکسیم کردن رابطه‌ی (۹) برآورد  $\bar{Y}_K$  به صورت زیر به دست می‌آید:

$$(۱۲) \quad \hat{\bar{Y}}_K = \frac{\sum_{q \in K} n_K^{(q)} \bar{y}_K^{(q)}}{\sum_{q \in K} n_K^{(q)}} = \frac{\sum_{q \in K} f^{(q)} \hat{Y}_K^{(q)}}{\sum_{q \in K} f^{(q)} \hat{N}_K^{(q)}}.$$

برای ماکسیم کردن رابطه‌ی (۱۰)،  $M$  را یک ماتریس  $d \times Q$  در نظر بگیرید که  $(i, j)$  امین درایه‌ی آن یک است اگر  $j \in K_i$  باشد و سایر جاها صفر است. همچنین فرض کنید که  $\hat{H}$  یک  $d \times Q$  به صورت زیر باشد، در این صورت:

$$\hat{H}_{i,j} = \begin{cases} \hat{N}_{K_i}^{(q)} & \text{اگر } q \in K_i \\ 0 & \text{در غیر این صورت} \end{cases}$$

که برای نمونه‌گیری تصادفی ساده،  $\hat{N}_K^{(q)} = (N^{(q)} / n^{(q)}) n_K^{(q)}$  است. قید (۱۱) به این معنی است که برداری که رابطه‌ی (۱۰) را ماکسیم می‌کند یعنی  $\hat{N}$  باید در  $M^T \hat{N} = \mathbf{h}$  صدق کند که در این رابطه  $\hat{N}$  حل معادله‌ی زیر است:

$$(۱۳) \quad \begin{bmatrix} (I - MM^+) (\text{diag} N)^{-1} \hat{H} \mathbf{f} \\ M^T \hat{N} - \mathbf{h} \end{bmatrix} = \begin{matrix} 0 \\ Q+d \end{matrix}$$

در معادله‌ی بالا  $\mathbf{h} = (N^{(1)}, \dots, N^{(Q)})^T$  و  $M^+$  وارون تعمیم‌یافته‌ی مور-پنرز ماتریس  $M$  است. اگر  $M$  وارون‌پذیر باشد آن‌گاه  $\hat{N} = M^{-T} \mathbf{h}$ . نظام معادله‌های (۱۳) نسبت به

$\hat{N}$  خطی نیست. لوهر و راثو [۱۰] با استفاده از قضیه‌ی نگاشت ضمنی یک تقریب خطی  $\tilde{N}$  برای  $\hat{N}$  یافتند. برآوردگر خطی  $\tilde{N}$  حل معادله‌ی زیر است:

$$(14) \quad \begin{bmatrix} (I - MM^+) (\text{diag} N)^{-1} (\text{diag} Mf) \\ M^T \end{bmatrix} \tilde{N} = \begin{bmatrix} (I - MM^+) (\text{diag} N)^{-1} \hat{H}f \\ h \end{bmatrix}$$

معادله‌ی (۱۴) برای یافتن واریانس مجانبی  $\hat{N}$  که با واریانس مجانبی  $\tilde{N}$  یکسان است مورد استفاده قرار می‌گیرد. همچنین این معادله برای محاسبه‌ی  $\hat{N}$  به کمک روش‌های تکراری به کار می‌رود. به این صورت که با در نظر گرفتن یک بردار از  $\hat{N}$ ، برآوردهای آغازین برای  $N$  و  $\tilde{N}_{K+1}$  حل معادله‌ی زیر باشد:

$$\begin{bmatrix} (I - MM^+) (\text{diag} \tilde{N}_K)^{-1} (\text{diag} Mf) \\ M^T \end{bmatrix} \tilde{N}_{K+1} = \begin{bmatrix} (I - MM^+) (\text{diag} \tilde{N}_K)^{-1} \hat{H}f \\ h \end{bmatrix}$$

به این ترتیب با استفاده از این معادله می‌توان اندازه‌ی جامعه‌ای،  $N$  را برآورد کرد. اگر برآورد آغازین  $\hat{N}$  قابل قبول باشد، آنگاه برآورد حاصل از این روش به سرعت به  $\hat{N}$  همگرا خواهد شد. در عمل مقدار آغازین را از متوسط  $\hat{N}_K^{(q)}$ ‌ها یا از چارچوب با بزرگ‌ترین اندازه‌ی نمونه در حوزه‌ی  $K$  محاسبه می‌کنند [۱۰].

### ۳-۳- روش‌های برآورد تک‌چارچوبی

روش‌های تک‌چارچوبی توسط بانکی‌یر [۱] و کالتون و اندرسون [۸] در موقعیت‌های دوچارچوبی پیشنهاد شدند. در این روش مشاهده‌ها داخل یک چارچوب تکی ترکیب و وزن‌ها در حوزه‌ی مشترک تعدیل می‌شوند. در این بخش روش تک‌چارچوبی به منظور برآورد مجموع جامعه‌ای با بیش از دو چارچوب به کار برده می‌شود. در برآوردگر بانکی‌یر [۱] لازم است که واحدها از نمونه‌های مختلف متمایز باشند. برای گسترش این برآوردگر به چارچوب‌های چندگانه، قرار دهید:

$$(15) \quad \tilde{w}_{i,S}^{(q)} = \sum_{K \subseteq F} \delta_i(K) \left( \sum_{j \in K} \frac{1}{w_i^{(j)}} \right)^{-1}$$

که  $w_i^{(j)}$  وزن واحد مشاهده‌ی  $i$ ام از چارچوب  $j$  است. با وزن تعدیل‌شده‌ی بالا برآوردگر تک‌چارچوبی برابر است با:

$$(۱۶) \quad \hat{Y}_S = \sum_{q=1}^Q \sum_{i \in S_q} \tilde{w}_{i,S}^{(q)} y_i$$

اگر هر نمونه خودوزن باشد (یعنی  $w_i^{(q)} = w_j^{(q)} = w^{(q)}$  برای همه  $i, j$  و  $q$ ‌ها)، آن‌گاه  $\hat{Y}_S$  می‌تواند به شکل (۳) نوشته شود. در این حالت با استفاده از فرمول‌های (۱)، (۱۵) و (۱۶) برآوردگر تک‌چارچوبی در آمارگیری‌های چندچارچوبی به صورت زیر محاسبه می‌شود:

$$\begin{aligned} \hat{Y}_S &= \sum_{q=1}^Q \sum_{K:q \in K} \left( w^{(q)} \sum_{j \in K} \frac{1}{w^{(j)}} \right)^{-1} \sum_{i \in S_q} w^{(q)} \delta_i(K) y_i \\ &= \sum_{q=1}^Q \theta_{q,S}^T \Gamma_q \hat{Y}_q \end{aligned}$$

با

$$\theta_{q,S}^T = \left[ \left( w^{(q)} \sum_{j \in K_1} \frac{1}{w^{(j)}} \right)^{-1}, \dots, \left( w^{(q)} \sum_{j \in K_d} \frac{1}{w^{(j)}} \right)^{-1} \right]$$

که  $\theta_{q,S}^T$  پارامتر برآوردگر تک‌چارچوبی یا همان وزن‌های تعدیل‌یافته برای چارچوب  $q$  است. هنگامی که اندازه‌ی چارچوب‌ها معلوم باشند برآوردگر تک‌چارچوبی می‌تواند با استفاده از برآورد نسبتی چنگک‌زنی [۱ و ۱۷]، یا برآورد رگرسیونی تعدیل شود. زمانی که همه‌ی آمارگیری‌های مجزا خودوزن باشند، برآوردگر تک‌چارچوبی برای هر تعدادی از چارچوب‌ها نیز محاسبه خواهد شد.

#### ۳-۴- برآوردگر تعدیل وزنی با سه یا بیش از سه چارچوب

در حالت کلی فرض کنید  $Q$  چارچوب  $A_1, \dots, A_Q$  وجود دارند. نمونه‌های احتمالاتی مستقل از چارچوب  $A_q$  با  $S_q$  برای  $q = 1, \dots, Q$  نشان داده می‌شود. واحد  $i$ ام در نمونه‌ی  $S_q$  دارای احتمال شمول  $\pi_i^{A_q}$  و وزن نمونه‌گیری  $w_i^{A_q}$  است. در مجموع  $D$  حوزه‌ی مجزا وجود دارد.

برآوردگر مجموع جامعه‌ای با استفاده از وزن‌های تعدیل‌یافته به شکل زیر است:

$$\hat{Y} = \sum_{i \in S_q} m_i^{A_q} w_i^{A_q} y_i$$

که در آن  $m_i^{Aq}$  عامل تعدیل وزن برای مشاهده‌ی  $i$ ام در  $S_q$  است. اگر حوزه‌ی  $d$  بخشی از چارچوب  $A_q$  نباشد،  $m^{(Aq,d)}$  (عامل تعدیل وزن برای هر چارچوب و حوزه) برابر با صفر است. البته تعدیل وزن مورد نظر باید در دو قید  $m^{(Aq,d)} \geq 0$  و  $\sum_{q=1}^Q m^{(Aq,d)} = 1$  برای  $d = 1, \dots, D$  صدق کند. بنا بر این هنگامی که مشاهده‌ی  $i$ ام از  $S_q$  در حوزه‌ی  $d$  است  $m_i^{Aq} = m_i^{(Aq,d)}$  [۱۲].

برای مثال این برآوردگر در آمارگیری‌های سه‌چارچوبی با استفاده از میانگین برآوردهای مجموع جامعه‌ای حوزه‌ها به صورت زیر محاسبه می‌شود:

$$\hat{Y}_{ave} = \hat{Y}_a^A + \hat{Y}_b^B + \hat{Y}_c^C + \frac{1}{4}(\hat{Y}_{ab}^A + \hat{Y}_{ab}^B) + \frac{1}{4}(\hat{Y}_{ac}^A + \hat{Y}_{ac}^C) + \frac{1}{4}(\hat{Y}_{bc}^B + \hat{Y}_{bc}^C) + \frac{1}{3}(\hat{Y}_{abc}^A + \hat{Y}_{abc}^B + \hat{Y}_{abc}^C)$$

این برآوردگر ساده‌ترین روش برای یافتن برآورد مجموع جامعه‌ای در آمارگیری‌های سه‌چارچوبی است که روش تعدیل وزنی نام دارد.

#### ۴- مطالعه‌ی شبیه‌سازی

در این بخش با استفاده از نرم‌افزار R، به روش شبیه‌سازی مونته کارلویی برآوردهای هارتلی، فولر-بورمیستر، ماکسیمم درست‌نمایی‌نما، تک‌چارچوبی و تعدیل وزنی برای مجموع جامعه‌ای در آمارگیری سه‌چارچوبی مقایسه می‌شوند. مشاهده‌های جامعه‌ای متناهی از آمارگیری سه‌چارچوبی برای حوزه‌ی  $K$ ام با استفاده از مدل احتمالاتی تصادفی زیر تولید شده است:

$$y_i = \mu_K + \varepsilon_i, \quad i = 1, \dots, N_K, \quad K = K_1, \dots, K_d$$

که  $\varepsilon_i$  دارای توزیع نرمال با میانگین صفر و واریانس یک است. میانگین حوزه‌ی  $K$ ام،  $\mu_K$  و اندازه‌ی جامعه‌ای هر حوزه با  $N_K$  نشان داده شده است. با اندازه‌ی جامعه‌ای معلوم

چارچوب‌ها و میانگین‌های مشخص، از هر سه چارچوب نمونه‌های تصادفی ساده و مستقل از هم گزینش شده‌اند.

برای انجام شبیه‌سازی، نمونه‌گیری را ۱۰۰۰ بار تکرار کرده و از هر نمونه ریشه‌ی دوم میانگین توان‌های دوم خطای تجربی برای برآوردهای مجموع جامعه‌ای محاسبه شده است. همچنین به منظور ارزیابی اثر اندازه‌ی نمونه‌ای بر میانگین توان دوم خطای برآوردها برای هر سه چارچوب اندازه‌های نمونه ۱۰۰، ۲۰۰ و ۵۰۰ در نظر گرفته شده است. نتیجه‌های شبیه‌سازی به ازای اندازه‌ی جامعه‌ای حوزه‌های مختلف و میانگین‌های متفاوت ارایه شده است.

یکی از معیارها برای بیان میزان درستی یک برآوردها، میانگین توان‌های دوم خطای تجربی است. با استفاده از شبیه‌سازی مونته کارلو، میانگین توان‌های دوم خطای تجربی برای تکرارهای مختلف نمونه‌گیری برابر است با:

$$EMSE = \frac{1}{B} \sum_{i=1}^B (\hat{t}_i - t)^2$$

که  $\hat{t}_i$  برآورد مجموع جامعه‌ای برای هر نمونه و  $t$  مجموع جامعه و معلوم است. ریشه‌ی دوم میانگین توان دوم خطای تجربی ( $\sqrt{EMSE}$ ) برآوردهای مجموع جامعه‌ای، در دو جامعه با اندازه‌ی جامعه‌ای حوزه‌ها و میانگین‌های متفاوت محاسبه و در زیر ارایه شده است. با توجه به شکل ۱، ترتیب حوزه‌ها در هر دو جامعه از چپ به راست به صورت  $\{1\}$ ،  $\{1,2\}$ ،  $\{2\}$ ،  $\{1,2,3\}$ ،  $\{1,3\}$ ،  $\{2,3\}$  و  $\{3\}$  است که اندازه‌ی جامعه‌ای در حوزه‌های  $\{1\}$ ،  $\{1,2\}$  و  $\{1,3\}$  در جامعه‌ی دوم افزایش داده شده است.

جامعه‌ی اول:

$$N^T = (2000, 2000, 2000, 2000, 2000, 2000, 2000), \mu^T = (11, 14, 12, 17, 16, 18, 19)$$

جامعه‌ی دوم:

$$N^T = (6000, 6000, 2000, 2000, 6000, 2000, 2000), \mu^T = (15, 10, 20, 17, 12, 18, 14)$$

با توجه به جدول ۱، برآوردها ماکسیم درست‌نمایی نما کم‌ترین  $\sqrt{EMSE}$  را در میان سایر روش‌ها دارد. برآوردها تک‌چارچوبی و روش تعدیل وزنی تقریباً با هم برابرند. با توجه به این‌که نمونه‌های تصادفی از هر یک از جامعه‌ی چارچوب‌ها انتخاب شده‌اند در نتیجه وزن‌های نمونه‌گیری برای واحدهای جامعه‌ی چارچوب‌ها درون هر حوزه یکسان است.

جدول ۱- ریشه‌ی دوم میانگین توان دوم خطای تجربی برآوردگرهای ۵گانه‌ی جامعه‌ی اول

SF	PML	FB	Hart	WA*	$n^{(3)}$	$n^{(2)}$	$n^{(1)}$
۵۴۶۰۸۵۳	۵۳۱۴۵۴۱	۵۴۶۰۸۰۷	۲۳۱۵۴۸۸	۵۴۶۰۸۵۳	۱۰۰	۱۰۰	۱۰۰
۳۸۶۸۴۱۹	۳۵۹۷۷۷	۳۸۶۸۳۵۹	۱۵۱۷۸۴۹	۳۸۶۸۴۱۹	۲۰۰	۲۰۰	۲۰۰
۲۳۹۳۸۵۴	۵۳۱۴۵۴۱	۲۳۹۳۶۲۳	۸۷۲۳۱۶۶	۲۳۹۳۸۵۴	۵۰۰	۵۰۰	۵۰۰

\*نمادهای WA, Hart, FB, PML و SF به‌ترتیب بیانگر برآوردگر تعدیل وزنی، هارتلی، فولر-بورمیستر، ماکسیمم درست‌نمایی‌نما و تک‌چارچوبی هستند.

از آنجایی که مقدار بهین  $\hat{\theta}_H$  در برآوردگر هارتلی به برآورد ماتریس کوواریانس مجموع جامعه‌ای برآوردشده بستگی دارد لذا این برآوردگر  $\sqrt{EMSE}$  بیش‌تری نسبت به سایر برآوردگرها دارد. اما برآوردگر فولر-بورمیستر با استفاده از اطلاعات اضافی (برآورد اندازه‌ی جامعه‌ای) تا حد زیادی این مشکل را مرتفع می‌سازد.

جدول ۲- ریشه‌ی دوم میانگین توان دوم خطای تجربی برآوردگرهای ۵گانه‌ی جامعه‌ی دوم

SF	PML	FB	Hart	WA*	$n^{(3)}$	$n^{(2)}$	$n^{(1)}$
۹۶۲۰۵۴۹	۸۹۶۴۶۷۸	۹۱۵۲۳۵۳	۲۶۲۲۵۰۲	۹۱۵۲۴۰۶	۱۰۰	۱۰۰	۱۰۰
۶۵۴۷۰۱۶	۶۲۶۲۰۷۵	۶۳۱۳۸۶۵	۱۸۴۳۰۱۶	۶۳۱۷۷۵	۲۰۰	۲۰۰	۲۰۰
۴۴۳۱۹۳۵	۴۲۲۵۰۶۷	۴۳۱۴۶۹۷	۱۶۴۹۰۹۷	۴۳۱۴۷۲۸	۵۰۰	۵۰۰	۵۰۰

همان‌طور که در جدول ۲ مشاهده می‌شود، روش ماکسیمم درست‌نمایی‌نما  $\sqrt{EMSE}$  کوچک‌تری برای برآورد مجموع جامعه‌ای داراست.

در عمل  $\sqrt{EMSE}$  برآوردگر فولر-بورمیستر بزرگ‌تر از  $\sqrt{EMSE}$  ماکسیمم درست‌نمایی‌نما است. برآوردگر فولر-بورمیستر به‌طور نظری بهین است اما برای محاسبه‌ی این برآوردگر یک ماتریس کوواریانس  $14 \times 14$  برای هر یک از سه نمونه باید برآورد شود و تغییرپذیری اضافی در برآورد کردن کوواریانس‌ها منجر به  $\sqrt{EMSE}$  بزرگ‌تری برای روش فولر-بورمیستر می‌شود.

برآوردگر هارتلی در آمارگیری سه‌چارچوبی عمل‌کرد ضعیفی را از خود نشان می‌دهد. در بعضی تکرارها، یک یا بیش‌تر مؤلفه‌های  $\hat{\theta}_H$  منفی بودند، برای مثال برآورد مجموع جامعه‌ای در حوزه‌ی  $abc$  به‌صورت  $\hat{Y}_{abc} = -0.5\hat{Y}_{abc}^A + 1.2\hat{Y}_{abc}^B + 0.3\hat{Y}_{abc}^C$  است. گاهی اوقات ناپایداری برآوردگر هارتلی سه‌چارچوبی با اندازه نمونه‌های کوچک موجب  $\sqrt{EMSE}$  بیش‌تری نسبت به برآوردگر میانگین‌گیری می‌شود. برای مثال، در جدول ۲،

$\sqrt{EMSE}$  برآوردگر هارتلی برای برآورد مجموع جامعه‌ای  $۲/۵۰۲۶۲۲۵$  است در حالی که  $\sqrt{EMSE}$  روش تعدیل وزنی  $۴۰۶/۹۱۵۲$  است. البته با افزایش اندازه‌ی نمونه‌ای در هر یک از سه چارچوب برآوردگر هارتلی پایدارتر خواهد شد. با این حال همچنان  $\sqrt{EMSE}$  بزرگ‌تری نسبت به روش ماکسیمم درست‌نمایی‌نما داشته است.

## ۵- بحث و نتیجه‌گیری

ما در این مقاله چندین برآوردگر نقطه‌ای از جمله تعدیل وزنی، هارتلی، فولر-بورمیستر، ماکسیمم درست‌نمایی‌نما و تک‌چارچوبی برای مجموع جامعه‌ای در آمارگیری‌های چندچارچوبی معرفی و درستی آن‌ها را با هم مقایسه کردیم.

برآوردگر ماکسیمم درست‌نمایی‌نما به دلیل این‌که از وزن‌های یکسانی برای همه‌ی متغیرها استفاده می‌کند نسبت به برآوردگرهای هارتلی و فولر-بورمیستر برتری دارد. دیگر برآوردگرهای معرفی شده در آمارگیری سه‌چارچوبی عمل‌کردشان به‌خوبی برآوردگر ماکسیمم درست‌نمایی‌نما نبود و به‌کارگیری آن‌ها در آمارگیری‌ها با بیش از دو چارچوب توصیه نمی‌شوند.

برآوردگرهای هارتلی و فولر-بورمیستر ناپایدار هستند، زیرا حل نظام معادله‌های آن‌ها به استفاده از ماتریس کوواریانس‌های برآوردشده‌ی بزرگ نیاز دارند.

برآوردگر تک‌چارچوبی نیز عمل‌کرد ضعیفی نسبت به برآوردگر ماکسیمم درست‌نمایی‌نما در بیش‌تر شبیه‌سازی‌ها داشته است. البته با استفاده از تعدیل وزن چنگک‌زنی یا رگرسیونی می‌توان به نتیجه بهتری دست یافت.

با این حال در همه‌ی شبیه‌سازی‌های صورت گرفته، برآوردگر ماکسیمم درست‌نمایی‌نما کارایی بالاتری نسبت به سایر برآوردگرها از خود نشان داده است. بنا بر این یک انتخاب مناسب برای برآورد مجموع در آمارگیری‌های سه‌چارچوبی خواهد بود.

در بسیاری از موارد، آمارگیری‌های چندچارچوبی بدون افزایش هزینه‌ی گردآوری

داده‌ها، درستی  $\frac{1}{(EMSE)}$  برآوردهای مقدارهای جامعه‌ای را بهبود می‌بخشند. البته در انتخاب برآوردگر و طرح نمونه‌گیری باید دقت کرد. برای مثال در جامعه‌های کمیاب با ترکیب چارچوب‌های ناقص، پوشش و درستی برآوردها افزایش می‌یابد.



## مرجع‌ها

- [1] Bankier, M.D. (1986). Estimator Based on Several Stratified Samples with Applications to Multiple Frame Surveys, *Journal of the American Statistical Association*, **81**, 1074-1079.
- [2] Fuller, W.A. and Burmeister, L.F. (1972). Estimators for Samples Selected from Two Overlapping Frames, in Proceeding of the Statistics Section, *American Statistical Association*, 245-249.
- [3] Haines, D.E. and Pollock, K.H. (1998). Combining Multiple Frames to Estimate Population Size and Totals, *Survey Methodology*, **24**, 79-88.
- [4] Hajek, J. (1971). Comment on an Essay on the Logical Foundations of Survey Sampling, Part one, In V. Godambe and D. Sprott (Eds.), *Foundations of Statistical Inference*. Holt, Rinehart and Winston Toronto.
- [5] Hartley, H.O. (1962). Multiple Frame Surveys, in Proceedings of the Social Statistics Section, *American Statistical Association*, 203-206.
- [6] Hartley, H.O. (1974). Multiple Frame Methodology and Selected Applications, *Sankhya*, Ser. C, **36**, 99-118.
- [7] Iachan, R. and Dennis, M.L. (1993). Multiple Frame Approach to Sampling the Homeless and Transient Population, *Journal of official Statistics*, **9**, 747-764.
- [8] Kalton, G. and Anderson, D.W. (1986). Sampling Rare Populations, *Journal of the Royal Statistical Society*, Ser. A, **149**, 65-82.
- [9] Lohr, S.L. and Rao, J.N.K. (2000). Inference in Dual Frame Surveys, *Journal of the American Statistical Association*, **95**, 271-280.
- [10] Lohr, S.L. and Rao, J.N.K. (2006). Estimation in Multiple-Frame Surveys, *Journal of the American Statistical Association*, **101**, 1019-1030.
- [11] Lohr, S.L. (2009). Multiple Frame Surveys, In *Handbook of Statistics, Sample Surveys: Design, Methods and Applications*, (Eds.,

- D. Pfeffermann and C. R. Rao). Amsterdam: North Holland, **29**, 71-88.
- [12] Lohr, S.L. (2011). Alternative Survey Sample Designs: Sampling with Multiple Overlapping Frames, *Survey Methodology*, **37**, 197-213.
- [13] Lohr, S.L. and Lu, Y. (2010). Gross Flow Estimation in Dual Frame Surveys, *Survey Methodology*, **36**, 13-22.
- [14] Mecatti, F. (2007). A Singl Frame Multiplicity Estimator for Multiple Frame Surveys, *Survey Methodology*, **33**, 151-157.
- [15] Rao, J.N.K. and Wu, C. (2010). Pseudo-Empirical Likelihood Inference for Dual Frame Surveys, *Journal of the American Statistical Association*, **105**, 1494-1503.
- [16] Skinner, C.J. (1991). On the Efficiency of Raking Ratio Estimation for Multiple Frame Surveys, *Journal of the American Statistical Association*, **86**, 779-784.
- [17] Skinner, C.J. and Rao, J.N.K. (1996). Estimation in Dual Frame Surveys With Complex Designs, *Journal of the American Statistical Association*, **91**, 349-356.

طیبه چگینی

فوق لیسانس آمار

تهران، خیابان شهید بهشتی، نبش احمد قصیر، دانشگاه علامه طباطبایی، دانشکده‌ی اقتصاد، گروه آمار.  
رایانشانی: taiebehchegini@yahoo.com

حمیدرضا نواب‌پور

دکتری آمار

تهران، خیابان شهید بهشتی، نبش احمد قصیر، دانشگاه علامه طباطبایی، دانشکده‌ی اقتصاد، گروه آمار.  
رایانشانی: h.navvabpour@src.ac.ir